THE JOURNAL OF BIOLOGICAL CHEMISTRY VOL. 289, NO. 6, pp. 3405–3415, February 7, 2014

Structural Insights into the Epimerization of β-1,4-Linked Oligosaccharides Catalyzed by Cellobiose 2-Epimerase, the Sole Enzyme Epimerizing Non-anomeric Hydroxyl Groups of Unmodified Sugars*

Received for publication, October 30, 2013, and in revised form, December 7, 2013 Published, JBC Papers in Press, December 20, 2013, DOI 10.1074/jbc.M113.531251

Takaaki Fujiwara[‡], **Wataru Saburi**[§], **Hirokazu Matsui**[§], **Haruhide Mori**[§], **and Min Yao**[‡1]

From the [‡]*Faculty of Advanced Life Science, Hokkaido University, Kita-10, Nishi-8, Kita-ku, Sapporo 060-0810 and the* [§]*Research Faculty of Agriculture, Hokkaido University, Kita-9, Nishi-9, Kita-ku, Sapporo 060-8589, Japan*

Background: The details of the catalytic mechanism of cellobiose 2-epimerase (CE) remains unclear.
Results: The crystal structures of *Rhodothermus marinus* CE in the apo form and complexes with its substrates/products 4-*O*-β-D-glucopyranosyl-D-mannnose, epilactose, or cellobiitol (reaction intermediate analog) were elucidated.
Conclusion: Epimerization catalyzed by CE proceeds through ring opening, deprotonation/reprotonation, carbon-carbon bond rotation, and ring closure.
Significance: This study yielded structural insights into epimerization catalyzed by CE.

Cellobiose 2-epimerase (CE) reversibly converts D-glucose residues into D-mannose residues at the reducing end of unmodified β1,4-linked oligosaccharides, including β-1,4-mannobiose, cellobiose, and lactose. CE is responsible for conversion of β1,4-mannobiose to 4-*O*-β-D-mannosyl-D-glucose in mannan metabolism. However, the detailed catalytic mechanism of CE is unclear due to the lack of structural data in complex with ligands. We determined the crystal structures of halothermophile *Rhodothermus marinus* CE (*Rm*CE) in complex with substrates/products or intermediate analogs, and its apo form. The structures in complex with the substrates/products indicated that the residues in the β5-β6 loop as well as those in the inner six helices form the catalytic site. Trp-322 and Trp-385 interact with reducing and non-reducing end parts of these ligands, respectively, by stacking interactions. The architecture of the catalytic site also provided insights into the mechanism of reversible epimerization. His-259 abstracts the H2 proton of the D-mannose residue at the reducing end, and consistently forms the *cis*-enediol intermediate by facilitated depolarization of the 2-OH group mediated by hydrogen bonding interaction with His-200. His-390 subsequently donates the proton to the C2 atom of the intermediate to form a D-glucose residue. The reverse reaction is mediated by these three histidines with the inverse roles of acid/base catalysts. The conformation of cellobiitol demonstrated that the deprotonation/reprotonation step is coupled with rotation of the C2-C3 bond of the open form of the ligand. Moreover, it is postulated that His-390 is closely related to ring opening/closure by transferring a proton between the O5 and O1 atoms of the ligand.

Cellobiose 2-epimerase (CE)[2] (EC 5.1.3.11), which was first identified in the ruminal anaerobic bacterium, *Ruminococcus albus* (1), catalyzes interconversion of the D-glucose residue into a D-mannose residue at the reducing end of β1,4-linked oligosaccharides, including β1,4-mannobiose, cellobiose, and lactose. Epimerases acting on carbohydrates and its derivatives can be divided into 25 groups (EC 5.1.3.-) depending on their catalytic reactions and substrate specificities. Most epimerases, converting the configuration of non-anomeric hydroxyl groups, act on modified active substrates harboring phosphate groups or nucleotide diphosphate groups. In contrast to most sugar epimerases, CE catalyzes the epimerization of unmodified sugars at the C2 position. Recently, the metabolic pathway of mannan involving CE was postulated for the CE-producing bacteria, *R. albus*, and *Bacteroides fragilis* (2, 3). During mannan metabolism, CE converts β1,4-mannobiose, produced by hydrolysis of mannan or phosphorolysis of β1,4-mannooligosaccharides catalyzed by β1,4-mannanase (EC 3.2.1.78) or β1,4-mannooligosaccharide phosphorylase (EC 2.4.1.-), respectively, into 4-*O*-β-D-mannosyl-D-glucose (Man-Glc). The resulting Man-Glc is consequently degraded into α-D-mannosyl phosphate and D-glucose by 4-*O*-β-D-mannosyl-d-glucose phosphorylase (EC 2.4.1.281). CE also converts lactose into the rare oligosaccharide, epilactose (4-*O*-β-D-galactosyl-D-mannnose) (4), which is useful for medical applications as a prebiotic agent. Epilactose stimulates bifidobacteria growth *in vivo*, suppresses the conversion of primary bile acid into secondary bile acid (5), and promotes absorption of some minerals (6, 7). Elucidation of the amino acid sequence of CE from *R. albus* (*Ra*CE) facilitated the identification of CEs in other bacteria (8). Among the CEs identified to date, *Rhodothermus marinus* CE (*Rm*CE) is highly stable at high temperatures, and converts lactose into epilactose with high efficiency in contrast to the other known CEs (9),

* This work was supported by Platform for Drug Discovery, Informatics, and Structural Life Science from the Ministry of Education, Culture, Sports, Science and Technology, Japan.

The atomic coordinates and structure factors (codes 3WKF, 3WKG, 3WKH, and 3WKI) have been deposited in the Protein Data Bank (http://wwpdb.org/).

[1] To whom correspondence should be addressed. E-mail: yao@castor.sci.hokudai.ac.jp.

[2] The abbreviations used are: CE, cellobiose 2-epimerase; Man-Glc, 4-*O*-β-D-mannosyl-D-glucose; AGE, *N*-acetylglucosamine 2-epimerase; PGI, phosphoglucose isomerases; Glc-Man, 4-*O*-β-D-glucosyl-D-mannose.

including *Ra*CE. Therefore, *Rm*CE is very attractive as an enzyme for production of epilactose.

Amino acid sequences of CEs show weak similarity to *N*-acetylglucosamine 2-epimerases (EC 5.1.3.8, AGEs) and aldose-ketose isomerases YihS. Recently, we determined the structure of *Ra*CE. The overall structure of *Ra*CE adopts an $(\alpha/\alpha)_6$ barrel similar to the catalytic domains of AGE and YihS, and therefore CE comprise the AGE superfamily together with AGE and YihS (10). The catalytic mechanism of CE was discussed based on structural comparison and mutational analysis of *Ra*CE (10). In CE, two essential histidine residues, corresponding to the putative general acid/base catalysts of AGEs, and a third histidine residue located at the bottom of the substrate binding site are required for catalysis. Epimerization catalyzed by CE is likely to proceed with the pair of histidine residues (*Ra*CE-His-243 and *Ra*CE-His-374) as general acid/base catalysts. Moreover, it appears that the roles of these two histidine residues as general acid/base catalysts may be reversed when the reverse reaction takes place. The histidine residues corresponding to *Ra*CE-His-243 and *Ra*CE-His-374 are completely conserved in enzymes belonging to the AGE superfamily, supporting the importance of these residues. These histidines were found in *Anabaena* sp. AGE (*a*AGE-His-239 and *a*AGE-His-382) (11) and *Salmonella enterica* YihS (*Se*YihS-His-248 and *Se*YihS-His-383) (12). In contrast, the function of the third histidine in *Ra*CE (*Ra*CE-His-184) conserved in CE and YihS was proposed (10). In the case of CE and YihS, the H2 proton of the unmodified sugars at the reducing end of substrates is not easily abstracted due to its high pK_a value. Therefore, the third histidine facilitates deprotonation of the C2 atom by interacting with the O2 atom of reducing end sugars for formation of the *cis*-enediol intermediate. Although the roles of the three most important histidine residues have been proposed, several key aspects of its catalytic mechanism remain unclear. First, structural insights into deprotonation and subsequent protonation of the chiral center the C2 atom, which proceed in the open form of the sugar, are unclear because there is a lack of structural information regarding enzymes belonging to the AGE superfamily in complex with the open form of the sugar. Second, it is unclear how CE forms *cis*-enediol or *cis*-enediolate intermediates via the ring opening step, followed by ring closure. Important residues responsible for ring opening/closure of the substrate were confirmed in *Se*YihS and phosphoglucose isomerase (PGI) from *Lactococcus lactis* (*Ll*PGI) (12, 13). Based on the structural analyses of *Ll*PGI and *Se*YihS, the nearby base catalyst (*Ll*PGI-Lys-518 and *Se*YihS-Glu-251) was suggested to abstract a proton from the 1-OH of the substrate, with the histidine residue (*Ll*PGI-His-388 and *Se*YihS-His-383) donating a proton to the O5 atom of the substrate as an acid catalyst. *Se*YihS-His-383 was, therefore, regarded as a bifunctional residue responsible for deprotonation of the O1 atom to form the *cis*-enediol intermediate as well as protonation of the O5 atom of D-mannose during ring opening (12). However, whether *Ra*CE-His-374 is involved in ring opening as well as *cis*-enediol intermediate formation is still controversial. Finally, the importance of a flexible loop close to the catalytic site of CE is controversial due to the lack of structures in complex with its substrates/products. In the case of *Se*YihS, a flexible loop (β11-α8 loop) was shown to participate in binding to the substrate, D-mannose. Indeed, the conformation of this loop was changed upon substrate binding, and Arg-238 and Phe-239 residing in this loop was brought inward to the catalytic site to interact with D-mannose (12). In the case of *Ra*CE, a flexible loop (β7-β8 loop) close to the catalytic site may compensate for the role of the β11-α8 loop in *Se*YihS (10). Thus, comparing the presence or absence of its substrate is required to solve the problems noted here.

In this study, we determined the three-dimensional structures of *Rm*CE in complex with its substrates/products, 4-*O*-β-D-glucosyl-D-mannose (*Rm*CE-Glc-Man) and epilactose (*Rm*CE-epilactose) as closed form ligands, and with the intermediate analog, cellobiitol (*Rm*CE-cellobiitol) as an open form ligand, as well as its apo-structure. These structures reveal the manner of interaction with ligands, and provide static snapshots of different states along the reaction pathway. Consequently, we propose the catalytic mechanism of this enzyme. In addition, we discuss important structural elements for high thermal stability of *Rm*CE with regard to the structural features and sequence comparison with other CEs.

EXPERIMENTAL PROCEDURES

Preparation of Recombinant RmCE—Recombinant *Rm*CE was produced in *Escherichia coli*, and *E. coli* proteins were removed by heat treatment as described previously (9). The samples thus obtained were further purified by anion exchange column chromatography using DEAE-Sepharose CL-6B (Amersham Biosciences) equilibrated with 50 mM Tris-HCl buffer (pH 8.5) to remove nucleic acid compounds. The flow-through fractions containing the samples were collected and dialyzed against 50 mM Tris-HCl buffer (pH 8.5) and 150 mM NaCl. Homogeneity of the samples was assessed by SDS-PAGE.

Crystallization and Data Collection—The purified *Rm*CE solubilized in 50 mM Tris-HCl buffer (pH 8.5) and 150 mM NaCl was concentrated at 7.5 mg ml^{-1} using Vivaspin-4 10K (GE Healthcare), and subsequently used for initial crystallization screening using a series of crystallization kits from Qiagen (Hilden, Germany) by the sitting-drop vapor-diffusion method. Briefly, 0.5 μl of protein solution was mixed with an equal volume of reservoir solution. Initial crystals of *Rm*CE were obtained with reservoir solution containing 0.1 M sodium acetate buffer (pH 4.5) and 1.0 M ammonium hydrogen phosphate. The initial conditions were then optimized to 0.1 M sodium acetate buffer (pH 4.5) and 1.2 M ammonium hydrogen phosphate using an aliquot of protein solution concentrated at 5.0 mg ml^{-1} in 30 mM Tris-HCl buffer (pH 8.5) and 60 mM NaCl. The crystals were grown to approximate dimensions of 0.1 × 0.1 × 0.05 mm within 5 days at 20 °C. For data collection, the crystals were soaked in cryoprotectant solution containing 20% (v/v) glycerol by repeated addition of 1.0 μl of cryoprotectant solution and discarding the 1.0-μl drop solution three times, followed by flash-cooling under a stream of liquid nitrogen at 100 K. Diffraction data were collected on beamline BL-41XU at SPring-8 (Hyogo, Japan). The asymmetric unit contained one molecule corresponding to a Matthews coefficient (14) of 1.83 $Å^3\ Da^{-1}$ and an estimated solvent content of 32.7%. To prepare the crystals of *Rm*CE complexed with ligands, a single crystal

TABLE 1
Data collection and refinement statistics

	*Rm*CE	*Rm*CE-Glc-Man	*Rm*CE-epilactose	*Rm*CE-cellobiitol
Data collection				
Wavelength (Å)	1.0000	1.0000	1.0000	1.0000
Space group	$P2_12_12_1$	$P2_12_12_1$	$P2_12_12_1$	$P2_12_12_1$
Unit cell parameters (Å)	$a = 41.9$	$a = 41.9$	$a = 41.9$	$a = 41.6$
	$b = 87.5$	$b = 87.7$	$b = 87.5$	$b = 87.0$
	$c = 94.0$	$c = 94.2$	$c = 94.0$	$c = 93.3$
Resolution (Å)	50.00–1.74 (1.81–1.74)	50.00–1.47 (1.56–1.47)	50.00–1.64 (1.74–1.64)	50.00–2.19 (2.31–2.19)
R_{merge} [a]	0.101 (0.448)	0.076 (0.511)	0.071 (0.486)	0.112 (0.586)
Redundancy	10.6 (9.1)	6.4 (6.5)	6.5 (6.4)	7.1 (7.0)
Completeness (%)	99.9 (99.8)	99.8 (99.2)	99.0 (93.6)	99.4 (96.6)
Number of unique reflections	36127 (3531)	59748 (9460)	42413 (6411)	18054 (2775)
$\langle I/\sigma(I)\rangle$	22.8 (4.0)	16.8 (3.8)	17.3 (3.6)	14.6 (3.8)
Refinement				
Resolution (Å)	32.0–1.74	39.8–1.47	43.8–1.64	38.1–2.19
R_{work}/R_{free}	0.158/0.185	0.167/0.194	0.160/0.194	0.161/0.224
Atoms				
Protein	3317	3345	3345	3317
Water	358	611	498	231
PO_4	10	5	5	5
Cl ion	1	1	1	1
Glc (ring form)				11
Glc (linear form)				12
Glc-Man		23		
Epilactose			23	
RMSD				
Bond length (Å)	0.006	0.006	0.006	0.007
Bond angle (°)	1.088	1.115	1.102	1.078
Ramachandran (%)				
Favored	98.8	98.8	98.8	98.8
Allowed	1.2	1.2	1.2	1.2
Outliers	0	0	0	0
Mean B factor (Å^2)				
Protein	13.94	12.13	18.73	23.37
Solvent	26.55	24.59	28.64	30.06

[a] $R_{merge} = \Sigma_{hkl}\Sigma_i|I_i(hkl) - \langle I_i(hkl)\rangle|/\Sigma_{hkl}\Sigma_i I_i(hkl)$, where i is the number of observations of a given reflection and $I(hkl)$ is the average intensity of the i observations. R_{free} was calculated with a 5% fraction of randomly selected reflections evaluated from refinement. The highest resolution shell is shown in parentheses.

was soaked in cryoprotectant solution containing 40% (w/v) sucrose and 10 mM cellobiose, 10 mM lactose, or 10 mM cellobiitol for 9 h at 20 °C. Diffraction data of each complex were collected on beamlines BL-1A and NW-12A at Photon Factory (Tsukuba, Japan). Data sets for *Rm*CE in complex with each ligand were processed using the *XDS* program suite (15), whereas data sets for *Rm*CE were processed using the *HKL2000* program suite (16). All data collection statistics are summarized in Table 1.

Structure Determination and Refinement—The structures of *Rm*CE and ligand-bound *Rm*CE were determined by the molecular replacement method with the program *AutoMR* in the *Phenix* program package (17). The structure of apo-*Rm*CE was solved using the structure of *Ra*CE (PDB code 3VW5) as a search model. Subsequently, ligand-bound *Rm*CE was determined using the structure of apo-*Rm*CE as a search model. Rotation and translation functions were calculated using data of 45.0 to 3.0-Å resolution. Several rounds of refinement were performed using the program *Phenix.refine* in the *Phenix* program suite, alternating with manual fitting and rebuilding based on $2F_o - F_c$ and $F_o - F_c$ electron densities using *COOT* (18). Then, water molecules, metal ions, monophosphates, substrates/products (Glc-Man and epilactose), and intermediate analog (cellobiitol) were built based on $2F_o - F_c$ and $F_o - F_c$ electron densities. The final refinement statistics and geometry defined by MolProbity (19) are summarized in Table 1.

RESULTS AND DISCUSSION

Overall Structures of RmCE in the Apo Form and in Complex with Ligands—The structure of the apo-*Rm*CE was determined by the molecular replacement method using the structure of *Ra*CE as a search model and refined to 1.74-Å resolution. As shown in the other epimerases/isomerases belonging to the AGE superfamily (10–12, 20), the overall fold of *Rm*CE was $(\alpha/\alpha)_6$ barrel (Fig. 1*A*). The final refined model was comprised of *Rm*CE monomer (residues Thr-3–Val-409), 358 water molecules, one chloride ion, and two phosphate ions. The chloride and phosphate ions would be derived from the buffers used for purification and crystallization of the recombinant *Rm*CE. The chloride ion was located in the vicinity of the catalytic site, and coordinated by the amino group of the peptide backbone of residues Gly-387, Tyr-389, and His-390 (Fig. 1*B*). One of the phosphate ions occupied the substrate binding site surrounded by the inner six helices of the $(\alpha/\alpha)_6$ barrel-fold, the other was bound to the positively charged patch formed by Arg-352, Arg-89, and Arg-96 in the symmetry-related molecule (Fig. 1, *C* and *D*). The latter phosphate ion, therefore, may affect crystal packing. Although cellobiose and lactose were used as substrates to obtain the complex structures by the crystal soaking method, electron densities of the epimerized β-anomeric Glc-Man and epilactose were clearly observed in the catalytic site (Fig. 2, *A–D*), suggesting that *Rm*CE turned over these disaccharides in the crystal. Although the activity of *Rm*CE is very low under the

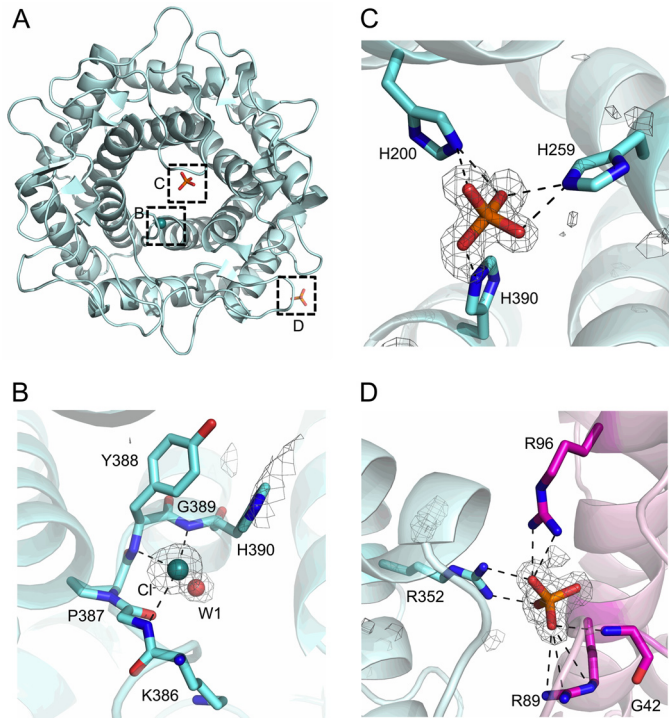

FIGURE 1. *A*, overall structure of apo-*Rm*CE displayed as a ribbon diagram. *B*, chloride ion binding site. The residues around the bound chloride ion are represented as *sticks*. Chloride ion and water molecule are represented as *spheres* in *red* and *green*, respectively. Polar interactions are shown by *dotted lines*. Electron density of the bound chloride ion and water in the omit $|F_o| - |F_c|$ map (*gray*) are calculated without the ligand and contoured at 3.0 σ. *C* and *D*, phosphate ion binding site. The residues around the bound phosphate ions are represented as *sticks*. Polar interactions are shown by *dotted lines*. Electron density of the bound phosphate in the omit $|F_o| - |F_c|$ map (*gray*) are calculated without the ligand and contoured at 3.0 σ.

conditions of the crystal soaking experiment (pH 4.5, 20 °C), a soaking time of 9 h would have been sufficient for epimerization to proceed. Indeed, it is well known that catalytic reactions of several proteins can be performed in the crystal (21, 22). In addition, selective binding to cellobiose or lactose was superior to that to sucrose (2-*O*-β-D-fructofuranosyl-α-D-glucoside) in the cryoprotectant as observed despite >100-fold greater concentration (40% (w/v) = 1.17 M) compared with cellobiose or lactose. The structures of *Rm*CE-Glc-Man, *Rm*CE-epilactose, and *Rm*CE-cellobiitol were determined and refined to 1.47-, 1.64-, and 2.19-Å resolutions, respectively. The residues Thr-3–Arg-412 of *Rm*CE-Glc-Man and *Rm*CE-epilactose were built, while residues Thr-3–Val-409 of *Rm*CE-cellobiitol were built. The chloride ion and the phosphate ions found in the molecular surface but not in the catalytic site were coordinated in the same fashion as observed in apo-*Rm*CE. The overall structures of *Rm*CE-Glc-Man, *Rm*CE-epilactose, and *Rm*CE-cellobiitol were very similar to that of apo-*Rm*CE with root mean square deviation values of 0.139 Å for 374 Cα atoms, 0.125 Å for 374 Cα atoms, and 0.149 Å for 380 Cα atoms, respectively. Thus, no noticeable rearrangement was induced upon ligand binding.

Binding Manner of Closed Form Ligands—In the structures of *Rm*CE in complex with the closed form ligands, several residues contribute to ligand recognition. The catalytic site was formed by residues in the β5-β6 loop, Ser-185, and Asp-188, as well as those in the inner six helices, Arg-66, Tyr-124, Asn-196, His-200, His-259, Glu-262, Tyr-307, Trp-322, Trp-385, and His-390 (Fig. 2, *B* and *D*). Aromatic residues are generally responsible for sugar recognition by stacking interaction (23). The indole groups of Trp-322 and Trp-385 were stacked onto pyranose rings at the reducing and non-reducing ends of the ligands, respectively. Trp-322 of *Rm*CE is completely conserved in enzymes belonging to the AGE superfamily, including AGEs and YihSs, whereas Trp-385 of *Rm*CE is only conserved in CEs. This observation is consistent with the results of mutational analyses indicating that these two tryptophan residues are important for ligand binding (10, 24). Moreover, the side chains of amino acid residues and water-mediated hydrogen bonds also recognize hydroxyl groups of ligands (Fig. 3*A*). Trp-322 recognizes the O2 atom of the D-mannose residue at the non-reducing end through an ordered water molecule. The amino groups of His-320 and Trp-322 help to position the water molecule regardless of ligand binding, suggesting that water is always trapped in the enzymes to orient the disaccharide properly in the catalytic site. Arg-66, which is completely conserved in the AGE superfamily and is essential for the catalytic activity (23), makes direct contact with the O6 atom of the D-mannose residue. Such an interaction was also observed in D-mannose-bound *Se*YihS (12). Extensive interactions were also identified around the reducing end sugar through Trp-265, Glu-326, and Arg-393, which are the other catalytically important residues confirmed by mutational analyses (23), and formed a salt bridge and hydrogen bond network to optimize the orientation and charge of His-390 and Glu-262 (Fig. 3*A*).

In addition to common interactions between two structures in complex with closed form ligands, the non-reducing end sugar mediates possible hydrogen bonding with Ser-185 and Asp-188 (Fig. 4*A*) in the β5-β6 loop positioned close to the active site. Ser-185 is completely conserved in known CEs characterized to date, whereas Asp-188 is diverse (Fig. 4*B*). In the *Rm*CE-Glc-Man structure, the O4 atom of the glucosyl residue at the non-reducing end directly forms hydrogen bonds with the side chain of Ser-185. In contrast, Ser-185 is slightly far from its O4 atom in the *Rm*CE-epilactose structure, which roughly correlates with high affinity toward cellobiose rather than lactose in most CEs except for *Dyadobacter fermentans* CE (*Df*CE), *Flavobacterium johnsoniae* CE (*Fj*CE), and *Pedobacter heparinus* CE (*Ph*CE) (8). On the other hand, diverse Asp-188 was positioned within hydrogen bonding distance to the O6 atom of both glucosyl and galactosyl residues of ligands. However, there is no relationship between the amino acid corresponding to Asp-188 and affinity toward cellobiose and lactose among CEs. One of the most striking aspects is that conformation of the β5-β6 loop in *Rm*CE is invariant independent of ligand binding, whereas the corresponding β11-α8 loop in *Se*YihS with a high degree of flexibility is visible in the catalytic site upon ligand binding (Fig. 4*C*).

Binding Manner of Open Form Ligand—The structure in complex with the reaction intermediate analog was required to understand the reaction mechanism of CE. Cellobiitol is an open chain form, and it could be an analog of the reaction intermediate. As shown in Fig. 2, *E* and *F*, this enzyme-intermediate analog complex structure is the first structural evidence for the

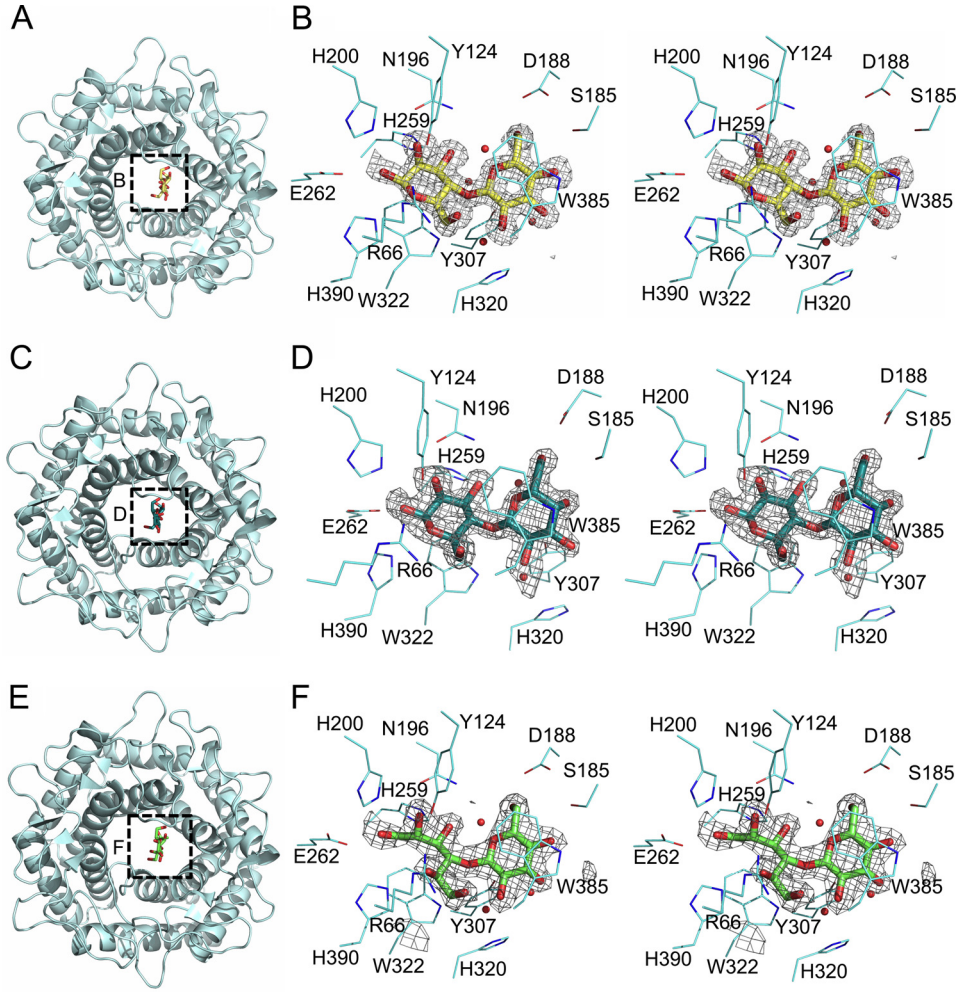

FIGURE 2. *A*, *C*, and *D*, the overall structures of *Rm*CE-Glc-Man, *Rm*CE-epilactose, and *Rm*CE-cellobiitol, respectively, displayed as ribbon diagrams. *B*, *D*, and *F*, stereo views of catalytic site architecture of *Rm*CE-Glc-Man, *Rm*CE-epilactose, and *Rm*CE-cellobiitol, respectively. The residues involved in substrate recognition are represented as *lines*. Water molecules are represented as *spheres* (*red*). Electron densities of the bound Glc-Man, epilactose, and cellobiitol in the omit $|F_o| - |F_c|$bar] map (*gray*) were calculated without the substrate and contoured at 3.0 σ.

active site among the enzymes belonging to the AGE superfamily. In the structure of *Rm*CE-cellobiitol, the pyranose ring of cellobiitol was recognized in an orientation similar to glucosyl and galactosyl residues at the non-reducing end of bound Glc-Man and epilactose, respectively (Figs. 2*F* and 3*B*). Moreover, the relative orientation of the D-glucitol part of cellobiitol was nearly identical to the reducing end sugar of Glc-Man and epilactose except for the positions of its C1 and O1 atoms (Fig. 5*A*). Projection of this D-glucitol part seen along with the C2-C3 bond indicated an eclipsed-like conformation with pseudoaxial orientation of the O2 atom (Fig. 5, *B* and *C*). The O2 atom of the D-glucitol part was turned around the C2-C3 bond about 90° from the equatorial position of the O2 atom of the ideal cellobiose. As a result, the C1-C2 bond of the D-glucitol part of cellobiitol was positioned between two His residues (His-259 and His-390) (Fig. 5*A*). These histidine residues correspond to *Ra*CE-His-243 and *Ra*CE-His-374, considered as acid/base catalysts during the epimerization step via the *cis*-enediol intermediate. On the basis of these observations, we propose that bound cellobiitol mimics the state just before abstraction of the H2 proton.

Structural Insights into the Ring Opening/Closure Step of the Catalytic Reaction with RmCE—The architectures of catalytic sites as well as the overall fold of *Rm*CE are similar to those of AGEs and YihSs, suggesting that these enzymes share a similar catalytic mechanism. The catalytic mechanism for CE was recently proposed based on sequence and structure comparison with AGEs and YihSs, as well as biochemical data (10, 23). Understanding the details of the reaction mechanism of CE, however, was restricted to the deprotonation/reprotonation process to form *cis*-enediol intermediates. Successfully determined crystal structures of *Rm*CE in complex with a series of ligands yielded further insights into interconversion of the D-glucose and D-mannose residues of disaccharide with *Rm*CE. Here, we provide the reaction mechanism proceeding via ring opening/closure and deprotonation/reprotonation coupled with rotation of the carbon-carbon bond to form the *cis*-enediol intermediate (Fig. 6).

Binding of cellobiitol to *Rm*CE reflects the requirement of ring opening to undergo epimerization. In addition, ring opening is a common process for the first step of sugar isomerization/epimerization, as exemplified by phosphoglucose isomer-

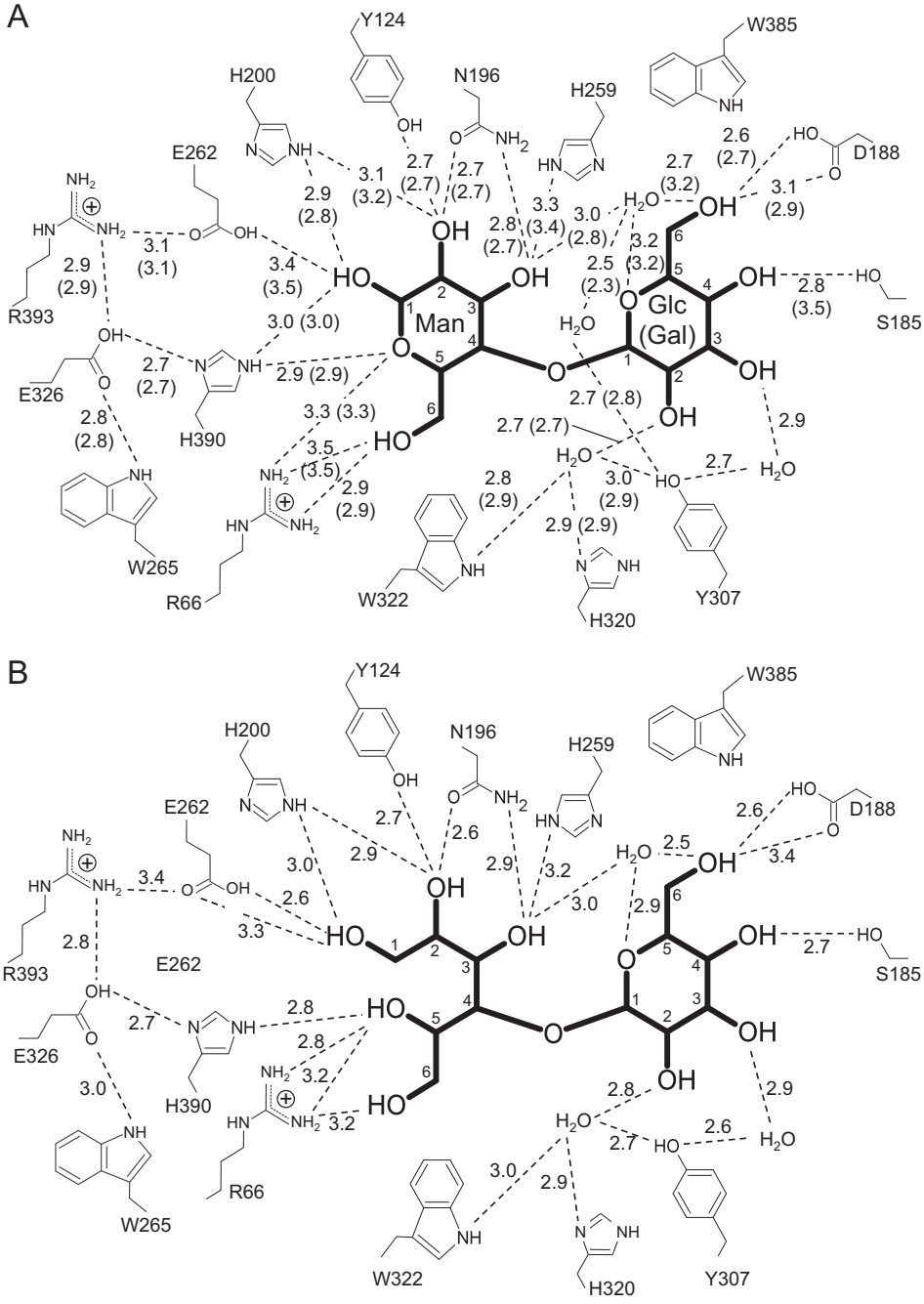

FIGURE 3. *A*, contact between Glc-Man or epilactose and *Rm*CE. The distances between specific atoms are shown in Å as *dotted lines*. *Numbers in parentheses* are for contact between epilactose and *Rm*CE. *B*, contact between cellobiitol and *Rm*CE. The distances between specific atoms are shown in Å as *dotted lines*.

ase, galactose mutarotase, and rhamnose isomerase (13, 25, 26) as well as *a*AGE and *Se*YihS (11, 12). The proposed mechanisms indicated that acid (A–H) and base ($B:^-$) catalysts are involved in ring opening/closure by donating a proton to the O5 atom of the reducing end moiety of the substrate and abstracting the proton from the O1 atom of the same moiety, respectively. As a consequence, the C1-O5 bond of the reducing end sugar was cleaved. The only candidate for the acid catalyst is His-390, as the basic residue Arg-66 also located within hydrogen bond distance to the O5 atom (3.3 Å) (Fig. 3*A*) is unlikely to act as an acid catalyst due to its high theoretical pK_a value, whereas those for the base catalysts are His-390, Glu-262, and His-200.

The Nϵ2 atom of His-390 was close to both O5 and O1 atoms of the reducing end of substrates, within 2.9 and 3.0 Å, respectively (Fig. 3*A*), which are reasonable for His-390 to serve as an acid/base catalyst for ring opening. In the direction of Glc → Man, initially deprotonated His-390 abstracts the proton from the O1 atom of the reducing end D-glucose residue, and then the resulting O5 anion accepts a proton from the protonated His-390, in agreement with the following role of His-390 as the base catalyst to form the *cis*-enediol intermediate (Fig. 7*A*). In the opposite direction of epimerization, initially protonated His-390 donates a proton to the O5 atom, and simultaneously His-390 abstracts the proton from the O1 atom, in agreement

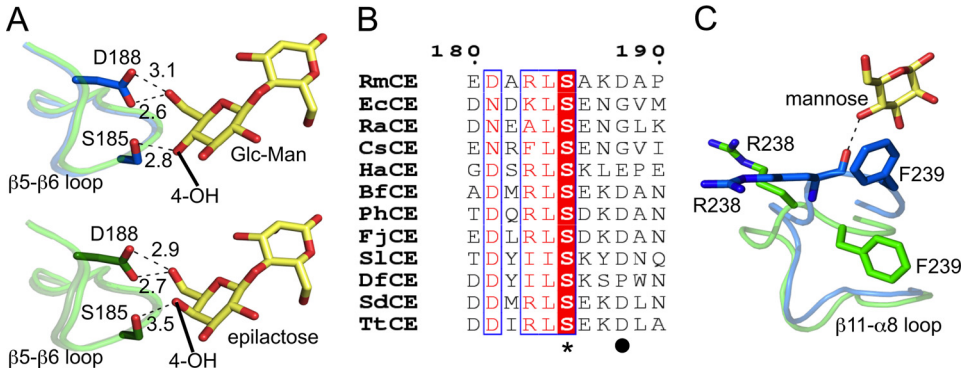

FIGURE 4. *A,* the structure of apo-form of *Rm*CE is superposed on that of Glc-Man-bound *Rm*CE (*upper panel*), and that of epilactose (*lower panel*). Apo structures in both panels are shown in *light green*. The β5-β6 loops in *Rm*CE in complex with Glc-Man (*blue*) and epilactose (*dark green*) are represented and the residues close to the substrate are depicted as *sticks*. *B,* multiple sequence alignment of several CEs. Serine and aspartic acid residues related to substrate binding are indicated by *stars* and *solid circles*, respectively. The multiple sequence alignment was performed using the programs ClustalW and ESPript. The sequences are as follows (GenBank™ accession numbers in parentheses): *Rm*CE (BAK61777.1), *Ra*CE (BAF81108.1), *Bf*CE (BAH23773.1), *Ec*CE (BAG68451.1), *Cs*CE (WP_011915904.1), and *Dt*CE (BAM66298.1). *C,* the ribbon diagrams of β11-α8 loop in *Se*YihS with or without D-mannose are shown in *green* and *blue*, respectively. The residues of the orientations of which are changed to make contact with the ligand are depicted by *sticks*.

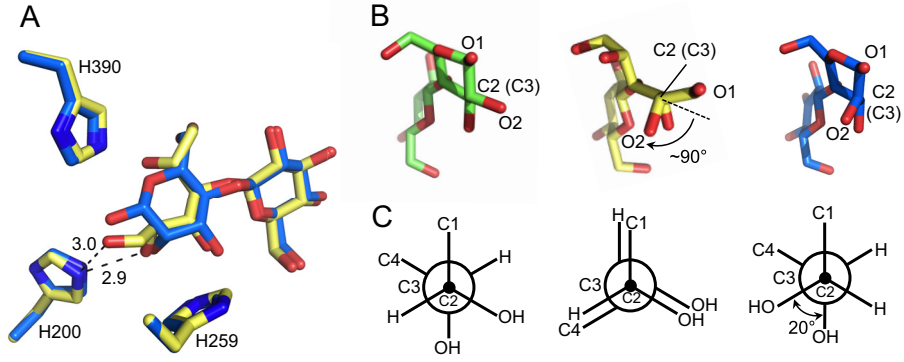

FIGURE 5. *A,* relative orientations of three catalytic histidine residues to cellobiitol and Glc-Man. Residues in *Rm*CE-cellobiitol and *Rm*CE-Glc-Man are represented as *sticks in yellow* and *blue*, respectively. *B*, the closed form sugars at the reducing end of cellobiose (*left*) and Glc-Man (*right*), and the sugar alcohol part of cellobiitol (*middle*) viewed along with the C2-C3 bond. *C,* projection of the closed form sugars at the reducing end of cellobiose (*left*) and Glc-Man (*right*), and the sugar alcohol part of cellobiitol (*middle*) viewed along with the C2-C3 bond.

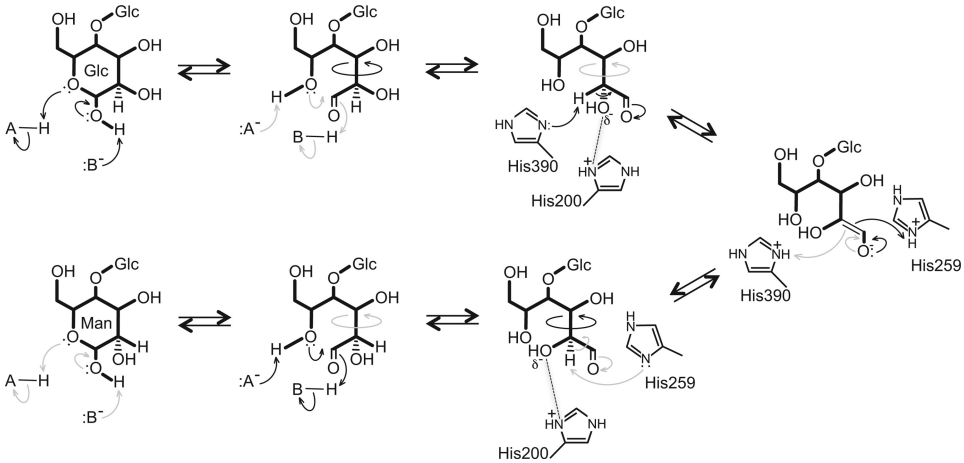

FIGURE 6. **Schematic diagram of the proposed epimerization catalyzed by *Rm*CE.** The electron transfer processes to convert glucose into mannose and its reverse conversion are represented by *black* and *gray arrows*, respectively.

with the following role of His-390 as the acid catalyst to donate a proton to the *cis*-enediol intermediate (Fig. 7*A*).

The mechanism of ring closure can be interpreted as the inverse reaction of ring opening. After the deprotonation/reprotonation step, rotation about the C2-C3 bond may occur again to bring the O1 atom of the reducing end moiety of the substrate close to the O5 atom of the same moiety (Fig. 6). Then, the electron pair on the O5 anion caused by deprotonation with His-390 attacks the electron-deficient C1 atom to remake the C1-O5 bond. Thus, this ring opening/closure mechanism involving only His-390 has no inconsistency for epimerization reactions in both directions.

On the other hand, it remains possible that Glu-262 or His-200 are related to ring opening/closure as a general base cata-

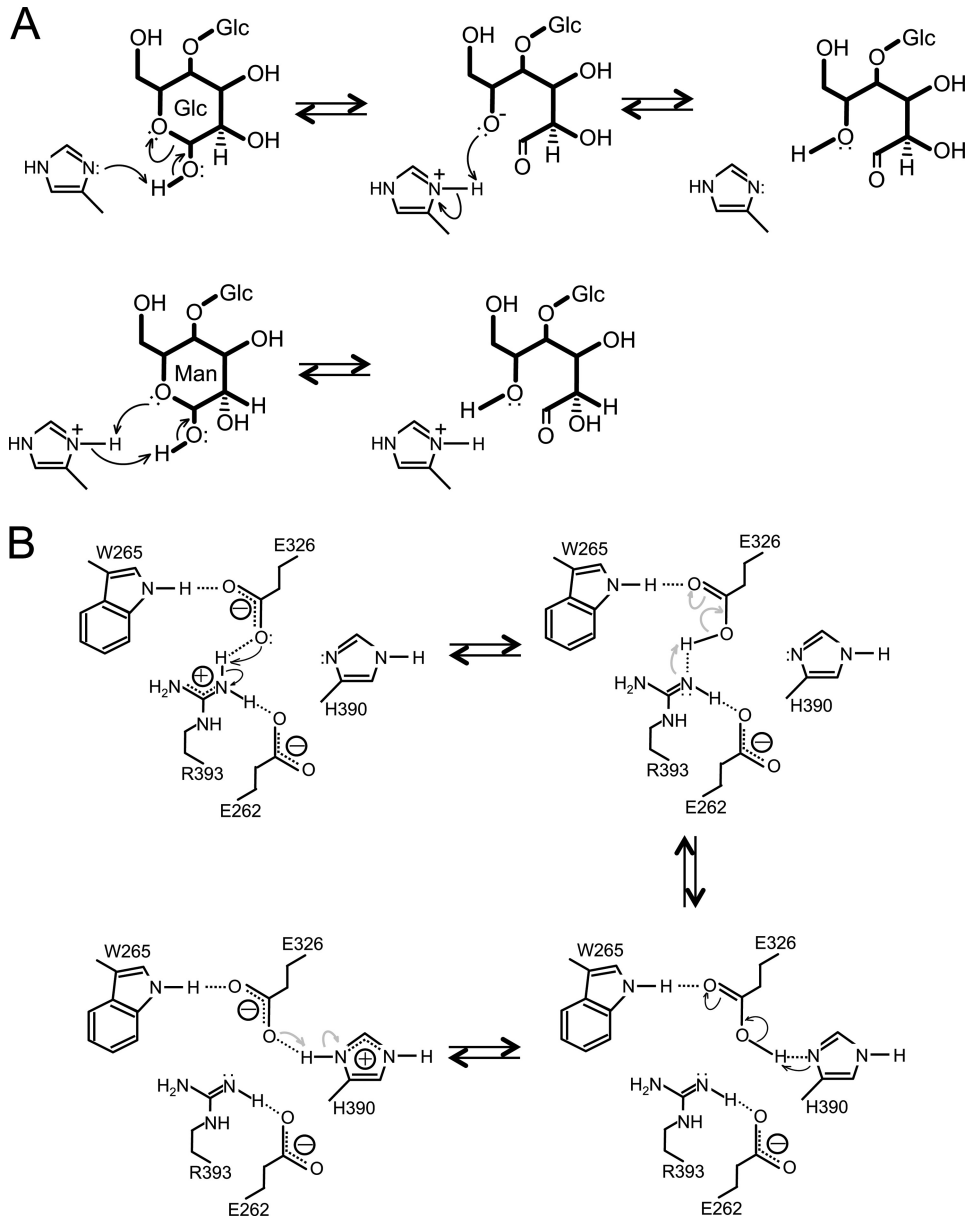

FIGURE 7. *A*, schematic diagram of the proposed ring opening mechanism involving His-390. The electron transfer reaction is represented by *arrows* (*black*). *B*, schematic diagram of the proposed proton-relay system. The electron transfer processes of forward and reverse reactions are represented by *black* and *gray arrows*, respectively.

lyst. However, the Oϵ1 atom of Glu-262 was slightly apart from the O1 atom within 3.5 Å (Fig. 3*A*), although it was proposed that the corresponding His-Glu pair in *Se*YihS (His-383–Glu-251) was involved in ring opening. His-200 is within 3.2 Å from the C1 atom of the reducing end of substrates (Fig. 3*A*), but this residue is not conserved in AGEs.

His-390 may be deprotonated when Glu-262 or His-200 acts as the base catalyst in ring opening. In the direction of Man → Glc, therefore, deprotonated His-390 should be protonated to donate a proton to the *cis*-enediol intermediate as the acid catalyst. When the open form of ligands is closed by the reverse reaction of ring opening, His-390 should be deprotonated to abstract the proton from the O5 atom of the reducing end part of the substrate as the base catalyst. In the direction of Glc → Man, therefore, His-390 protonated by abstracting a proton to form the *cis*-enediol intermediate should be deprotonated to act as the base catalyst for sequential ring closure. If His-390 acts as the acid catalyst for ring opening, therefore, there must be a system to remove the inconsistency on the charged state of His-390. A possible mechanism for exchanging the charged states of His-390 is that a proton may be relayed via nearby Glu-326, and Arg-393, forming a polar interaction network together with His-390 (Fig. 7*B*).

Structural Insights into the Deprotonation/Reprotonation Step of the Catalytic Reaction with RmCE—As mentioned previously (27), the H2 proton of the reducing end part of the substrate should be abstracted to form a *cis*-enediol intermediate after ring opening. The binding structure of the open form of the intermediate analog cellobiitol allowed us to deduce the detailed mechanism of epimerization by CE. Considering the

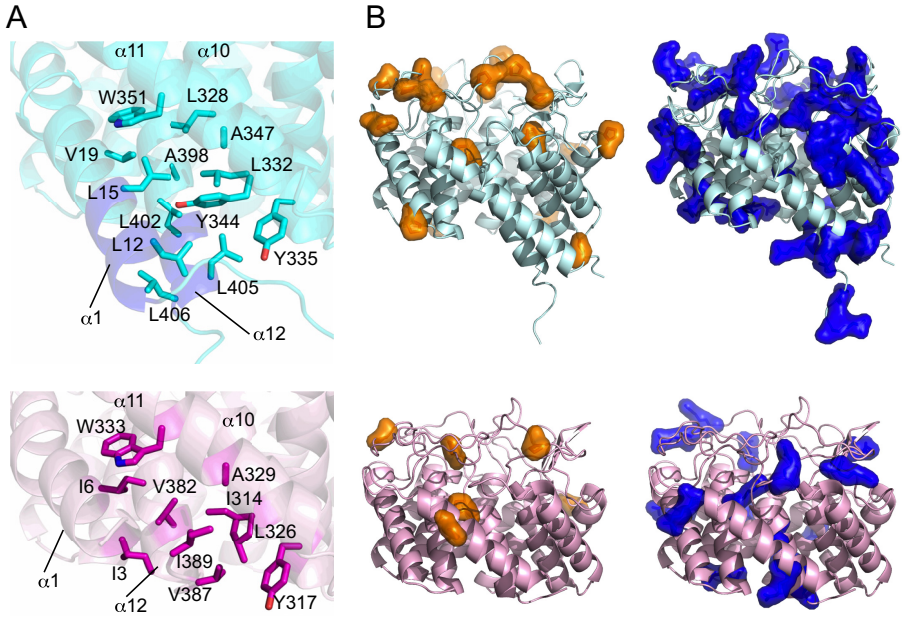

FIGURE 8. **Relationships between structural features to thermal stability of *Rm*CE.** The structure displayed is the Glc-Man-bound form of *Rm*CE. *A*, hydrophobic core formed by helices α1, α10, α11, and α12. Ribbon diagrams of *Rm*CE (*upper panel*) and *Ra*CE (*lower panel*) are shown, and hydrophobic residues are represented as *sticks*. Extended parts of helices in *Rm*CE are shown in *dark blue*. *B,* positions of Pro and Arg residues in *Rm*CE (*upper left* and *upper right*, respectively) and *Ra*CE (*lower left* and *lower right*, respectively) are indicated. Surface representations of Pro and Arg residues are shown in *orange* and *blue*, respectively.

stereochemistry of cellobiitol (Fig. 5*A*), the H2 proton of the D-glucitol part is close to His-390 by rotation about the C2-C3 bond, whereas it is very difficult for His-390 to approach the same H2 proton in the structure of *Rm*CE in complex with closed form ligands. When converting the D-glucose residue at the reducing end of substrates to the D-mannose residue, His-390 is the most feasible candidate to abstract the H2 proton from the reducing end D-glucose residue after rotation about the C2-C3 bond. When the proton is supplied from the opposite side of the general base catalyst His-390 across the C1=C2 double bond in the state of *cis*-enediol intermediate, the configuration around the stereogenic center of C2 converts. The orientation of the C1=C2 double bond that results in rotation about the C2-C3 bond makes it possible that the C2 atom accepts a proton from the protonated His-259 as a general acid catalyst. In this situation, His-200, corresponding to *Ra*CE-His-184, maintained in close proximity to the 2-OH group of the D-glucitol part by about 2.9 Å, fulfills the proposed roles of *Ra*CE-His-184 as assistant to depolarization of the 2-OH group. The distance between the Nϵ2 atom of His-200 and the O1 atom of the D-glucitol part of cellobiitol was also about 3.0 Å, suggesting that the negatively charged enolate anion of the actual intermediate is stabilized by the positive charge on His-200 (Fig. 5*A*).

On the other hand, the H2 proton of the D-mannose residue at the reducing end was close to His-259 in the structure of *Rm*CE in complex with the closed form ligands. When converting this D-mannose residue to D-glucose residue, therefore, His-259 is likely to abstract the H2 proton from the D-mannose residue independent of rotation about the C2-C3 bond. Subsequently, to convert the configuration of the C2 atom, the C2 atom accepts a proton from the opposite side of His-259 across to the C1=C2 double bond, where His-390 is located. Hence, His-390 is a reasonable candidate to function as an acid catalyst in this direction, and agreement with their roles proposed as reversible general acid/base catalysts in epimerization. Taken together, our results strongly suggest the requirement of rotation about the C2-C3 bond coupled with deprotonation/reprotonation to achieve reversible epimerization associated with three His residues (His-200, His-259, and His-390).

Taken together, the structural studies of *Rm*CE presented here provide molecular insights into the reaction mechanism for CE partly postulated based on structure analysis of the apo-form of *Ra*CE. Further studies, such as neutron diffraction analyses exemplified by xylose isomerase (28) and ultra-high resolution crystal structure analysis, would clarify the proton transfer and charged states of catalytic residues in catalysis by visualizing hydrogen atoms.

Relationship between Thermostability and Structure in RmCE—The overall fold of *Rm*CE was similar to that of *Ra*CE with a root mean square deviation of 0.88 Å for 338 Cα atoms. Although the conformations of both enzymes are almost identical, their α1 and α12 helices have different lengths; the α1 and α12 helices of *Rm*CE are longer than those of *Ra*CE by 1.5 turns each (Fig. 8*A*). Especially, five Leu residues in this extended 1.5 turns of helices α1 and α12 (Leu-12, Leu-15, Leu-402, Leu-405, and Leu-406) participated in forming a hydrophobic core with other hydrophobic residues, Val-19, Leu-328, Leu-332, Tyr-335, Tyr-344, Ala-347, Trp-351, and Ala-398 (Fig. 8*A*). The hydrophobic interactions observed in the buried surface area generally contribute to entropic stabilization of proteins (29). In addition, the contents of Pro and Arg residues in *Rm*CE are much higher than those in other CEs (Table 2), although *Caldicellulosiruptor saccharolyticus* CE (*Cs*CE) and *Dictyoglomus turgidum* CE (*Dt*CE) have less Pro and Arg residues regardless of their high optimum temperatures. Twenty Pro residues (1.7–

TABLE 2
Relationships between the compositions of prolines and arginines and optimum temperature in several CEs

	Proline	Arginine	Optimum temperature (° *C*)
*Rm*CE	20 (4.9%)	46 (11.2%)	80
*Cs*CE	4 (1.0%)	15 (3.8%)	75
*Dt*CE	11 (2.8%)	12 (3.1%)	70
*Df*CE	12 (3.1%)	16 (4.1%)	50
*Bf*CE	8 (2.0%)	25 (6.4%)	45
*Sl*CE	18 (4.5%)	18 (4.5%)	45
*Ha*CE	14 (3.4%)	21 (5.0%)	45
*Tt*CE	10 (2.4%)	23 (5.4%)	35
*Ec*CE	9 (2.2%)	18 (4.4%)	35
*Sd*CE	8 (2.0%)	16 (3.9%)	35
*Ph*CE	7 (1.8%)	14 (3.5%)	35
*Fj*CE	8 (2.0%)	7 (1.8%)	35
*Ra*CE	6 (1.5%)	16 (4.1%)	30

5.0 times more than the number of these residues in other CEs described above) were situated in the loops or the terminals of helices of *Rm*CE, and contribute to restricting the flexibility of the peptide backbone (Fig. 8*B*). It is assumed that reduced flexibility of unstructured regions is one of the reasons for the acquisition of thermostability by this enzyme (30). The content of the other key residue, Arg, in *Rm*CE is also 1.8–6.6 times higher than in the other CEs. The higher frequency of Arg residues in the exposed state is one of the critical factors related to protein thermal stability (Fig. 8*B*) (31). This feature has been explained by the observation that Arg residues participate in formation of pairs with negatively charged amino acid residues, such as Asp and Glu residues, on the molecular surface. As a result, hydrophobic interactions in the buried surface area and hydrophilic interactions on the molecular surface may contribute to the thermal stability of *Rm*CE.

Acknowledgments—We thank the staff of beamline BL-41XU at SPring-8, and BL-1A and NW-12A at the Photon Factory for assistance with data collection.

REFERENCES

1. Tyler, T. R., and Leatherwood, J. M. (1967) Epimerization of disaccharides by enzyme preparations from *Ruminococcus albus*. *Arch. Biochem. Biophys.* **119,** 363–367
2. Senoura, T., Ito, S., Taguchi, H., Higa, M., Hamada, S., Matsui, H., Ozawa, T., Jin, S., Watanabe, J., Wasaki, J., and Ito, S. (2011) New microbial mannan catabolic pathway that involves a novel mannosylglucose phosphorylase. *Biochem. Biophys. Res. Commun.* **408,** 701–706
3. Kawahara, R., Saburi, W., Odaka, R., Taguchi, H., Ito, S., Mori, H., and Matsui, H. (2012) Metabolic mechanism of mannan in ruminal bacterium, *Ruminococcuss albus*, involving two mannoside phosphorylases and cellobiose 2-epimerase. Discovery of a new carbohydrate phosphorylase, β-1,4-mannooligosaccharide phosphorylase. *J. Biol. Chem.* **287,** 42389–42399
4. Ito, S., Taguchi, H., Hamada, S., Kawauchi, S., Ito, H., Senoura, T., Watanabe, J., Nishimukai, M., Ito, S., and Matsui, H. (2008) Enzymatic properties of cellobiose 2-epimerase from *Ruminococcus albus* and the synthesis of rare oligosaccharides by the enzyme. *Appl. Microbiol. Biotechnol.* **79,** 433–441
5. Watanabe, J., Nishimukai, M., Taguchi, H., Senoura, T., Hamada, S., Matsui, H., Yamamoto, T., Wasaki, J., Hara, H., and Ito, S. (2008) Prebiotic properties of epilactose. *J. Dairy Sci.* **91,** 4518–4526
6. Nishimukai, M., Watanabe, J., Taguchi, H., Senoura, T., Hamada, S., Matsui, H., Yamamoto, T., Wasaki, J., Hara, H., and Ito S. (2008) Effects of epilactose on calcium absorption and serum lipid metabolism in rats. *J. Agric. Food. Chem.* **56,** 10340–10345
7. Suzuki, T., Nishimukai, M., Takechi, M., Taguchi, H., Hamada, S., Yokota, A., Ito, S., Hara, H., and Matsui, H. (2010) The nondigestible disaccharide epilactose increases paracellular Ca absorption via rho-associated kinase- and myosin light chain kinase-dependent mechanisms in rat small intestines. *J. Agric. Food Chem.* **58,** 1927–1932
8. Ojima, T., Saburi, W., Yamamoto, T., Mori, H., and Matsui, H. (2013) Identification and characterization of cellobiose 2-epimerases from various aerobes. *Biosci. Biotechnol. Biochem.* **77,** 189–193
9. Ojima, T., Saburi, W., Sato, H., Yamamoto, T., Mori, H., and Matsui, H. (2011) Biochemical characterization of a thermophilic cellobiose 2-epimerase from a thermohalophilic bacterium, *Rhodothermus marinus* JCM9785. *Biosci. Biotechnol. Biochem.* **75,** 2162–2168
10. Fujiwara, T., Saburi, W., Inoue, S., Mori, H., Matsui, H., Tanaka, I., and Yao, M. (2013) Crystal structure of *Ruminococcus albus* cellobiose 2-epimerase. Structural insights into epimerization of unmodified sugar. *FEBS Lett.* **587,** 840–846
11. Lee, Y. C., Wu, H. M., Chang, Y. N., Wang, W. C., and Hsu, W. H. (2007) The central cavity from the $(\alpha/\alpha)_6$ barrel structure of Anabaena sp. CH1 *N*-acetyl-D-glucosamine 2-epimerase contains two key residues for reversible conservation. *J. Mol. Biol.* **367,** 895–908
12. Itoh, T., Mikami, B., Hashimoto, W., and Murata, K. (2008) Crystal structure of YihS in complex with D-mannose. Structural annotation of *Escherichia coli* and *Salmonella enterica* yihS-encoded proteins to an aldose-ketose isomerases. *J. Mol. Biol.* **377,** 1443–1459
13. Solomons, J. T., Zimmerly, E. M., Burns, S., Krishnamurthy, N., Swan, M. K., Krings, S., Muirhead, H., Chirgwin, J., and Davies, C. (2004) The crystal structure of mouse phosphoglucose isomerases at 1.6 Å resolution and its complex with glucose 6-phosphate reveals the catalytic mechanism of sugar ring opening. *J. Mol. Biol.* **342,** 847–860
14. Kantardjieff, K. A., and Rupp, B. (2003) Matthewes coefficient probabilities. improved estimates for unit cell contents of proteins, DNA, and protein-nucleic acid complex crystals. *Protein Sci.* **12,** 1865–1871
15. Kabsch, W. (2010) *XDS. Acta Crystallogr. Sect. D* **66,** 125–132
16. Otwinowski, Z., and Minor, W. (1997) Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol.* **276,** 307–326
17. Adams, P. D., Afonine, P. V., Bunkoczi, G., Chen, V. B., Davis, I. W., Echols, N., Headd, J. J., Hung, L. W., Kapral, G. J., Grosse-Kunstleve, R. W., McCoy, A. J., Moriarty, N. W., Oeffner, R., Read, R. J., Richardson, D. C., Richardson, J. S., Terwilliger, T. C., and Zwart, P. H. (2010) PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr. Sect. D* **66,** 213–221
18. Emsley, P., and Cowtan, K. (2004) COOT. Model-building tools for molecular graphic. *Acta Crystallogr. D Biol. Crystallogr.* **60,** 2126–2132
19. Chen, V. B., Arendall, W. B.. 3rd, Headd, J. J., Keedy, D. A., Immormino, R. M., Kapral, G. J., Murray, L. W., Richardson, J. S., Richardson, D. C. (2010) MolProbity. All-atom structure validation for macromolecular crystallography. *Acta Crystallogr. D Biol. Crystallogr.* **66,** 12–21
20. Itoh, T., Mikami, B., Maru, I., Ohta, Y., Hashimoto, W., and Murata, K. (2000) Crystal structure of *N*-acetyl-D-glucosamine 2-epiemrase from porcine kidney at 2.0 Å resolution. *J. Mol. Biol.* **303,** 733–744
21. Mozzarelli, A., and Rossi, G. L. (1996) Protein function in the crystal. *Annu. Rev. Biophys. Biomol. Struct.* **25,** 343–365
22. Yamane, J., Ohyabu, N., Yao, M., Takemoto, H., and Tanaka, I. (2010) In-crystal chemical ligation for lead compound generation. *J. Appl. Crystallogr.* **43,** 1329–1337
23. Asensio, J. L., Ardá, A., Cañada, F. J., and Jiménez-Barbero, J. (2013) Carbohydrate-aromatic interactions. *Acc. Chem. Res.* **46,** 946–954
24. Ito, S., Hamada, S., Ito, H., Matsui, H., Ozawa, T., Taguchi, H., and Ito, S. (2009) Site-directed mutagenesis of possible catalytic residues of cellobiose 2-epimerase from *Ruminococcus albus*. *Biotechnol. Lett.* **31,** 1065–1071
25. Thoden, J. B., Kim, J., Raushel, F. M., and Holden, H. M. (2003) The catalytic mechanism of galactose mutarotase. *Protein Sci.* **12,** 1051–1059
26. Yoshida, H., Yoshihara, A., Teraoka, M., Yamashita, S., Izumori, K., and Kamitori, S. (2012) Structure of L-rhamnose isomerase in complex with L-rhamnopyranose demonstrates the sugar-ring opening mechanism and the role of a substrate sub-binding site. *FEBS Open. Biol.* **7,** 35–40

27. Amein, M., and Leatherwood, J. M. (1969) Mechanism of cellobiose 2-epimerase. *Biochem. Biophys. Res. Commun.* **36,** 223–227
28. Kovalevsky, A. Y., Hanson, L., Fisher, S. Z., Mustyakimov, M., Mason, S. A., Forsyth, V. T., Blakeley, M. P., Keen, D. A., Wagner, T., Carrell, H. L., Katz, A. K., Glusker, J. P., and Langan, P. (2010) Metal ion roles and the movement of hydrogen during reaction catalyzed by D-xylose isomerases. A joint X-ray and neutron diffraction study. *Structure* **18,** 688–699
29. Kellis, J. T. Jr., Nyberg, K., Sali, D., and Fersht, A. R. (1988) Contribution of hydrophobic interactions to protein stability. *Nature* **333,** 784–786
30. Takano, K., Higashi, R., Okada, J., Mukaiyama, A., Tadokoro, T., Koga, Y., and Kanaya, S. (2009) Proline effect on thermostability and slow unfolding of a hyperthermophilic protein. *J. Biochem.* **145,** 79–85
31. Pack, S. P., and Yoo, Y. J. (2004) Protein thermostability. Structure-based difference of amino acid between thermophilic and mesophilic proteins. *J. Biotechnol.* **111,** 269–277

Structural Insights into the Epimerization of β-1,4-Linked Oligosaccharides Catalyzed by Cellobiose 2-Epimerase, the Sole Enzyme Epimerizing Non-anomeric Hydroxyl Groups of Unmodified Sugars

Takaaki Fujiwara, Wataru Saburi, Hirokazu Matsui, Haruhide Mori and Min Yao

J. Biol. Chem. 2014, 289:3405-3415.
doi: 10.1074/jbc.M113.531251 originally published online December 20, 2013